

Dynamic simulations of water at constant chemical potential

Jie Ji, Tahir Çağın,^{a)} and B. Montgomery Pettitt
Chemistry Department, University of Houston, Houston, Texas 77204-5641

(Received 20 June 1991; accepted 7 October 1991)

The grand molecular dynamics (GMD) method has been extended and applied to examine the density dependence of the chemical potential of a three-site water model. The method couples a classical system to a chemical potential reservoir of particles via an ansatz Lagrangian. Equilibrium properties such as structure and thermodynamics, as well as dynamic properties such as time correlations and diffusion constants, in open systems at a constant chemical potential, are preserved with this method. The average number of molecules converges in a reasonable amount of computational effort and provides a way to estimate the chemical potential of a given model force field.

I. INTRODUCTION

The utility of molecular dynamics simulations lies in the construction of a particular ensemble in which macroscopic quantities may be obtained directly through time averages of microscopic variables. Many phenomena of chemical interest occur in open systems where matter and energy are exchanged with the surroundings.^{1,2} In addition, there is considerable interest in calculating the chemical potentials which govern open systems.²⁻⁹ Thus, it is of interest to consider other ensembles such as the grand canonical or grand adiabatic for utilization in dynamic computer simulations of condensed phase systems. Until recently, the ensembles studied in molecular dynamics simulations have been confined to those containing a fixed number of particles.¹⁰⁻¹⁴ This was due to the theoretical and technical difficulties associated with the discrete nature of the physical particle number vs the continuous equations of motion of the particles.¹⁵⁻¹⁷ However, in general, systems of interest such as realistic chemical and biochemical assemblies of molecules usually consist of several atomic or molecular species in which the number of molecules (the concentrations of the species) may fluctuate by material exchange with the surroundings at constant chemical potential.¹⁸

Molecular dynamics studies of free energies and chemical potentials have been achieved by using thermodynamic integration, or perturbation methods to obtain approximate free energy differences between two similar systems with a different number of particles, say N and $N + 1$, or different temperatures.⁴⁻⁹ These studies are computationally quite intensive for molecular liquids since the convergence in the measured quantities or the integrands is slow (in excess of several hundred picoseconds) and often are nonmonotonic as a function of the simulated time.^{19,20} By switching to the proper statistical mechanical ensemble, the free energy or chemical potential problem can be posed in a very different fashion. For simulations in the grand ensembles, by fixing the chemical potential, the average number/concentrations can then be measured. Not only can one map the chemical

potential vs density/concentration relation, but the chemical potential, at a target density, can be determined by interpolation.

We have recently proposed a new dynamic computer simulation method which provides dynamics for the number of molecules where the number of particles is taken to be a continuous variable; in essence, a fractional molecule was introduced in addition to the normal whole molecules.²¹ The proposed equations of motion, based on a coupling parameter approach, have been used to consider simple fluid systems with a variable number of molecules.²² At equilibrium, the average number of particles \bar{N} is well-defined and driven by the chemical potential μ .^{17,23,24} In this method, the number of molecules emerges as a continuous dynamic variable and, therefore, can have any real value during the dynamics simulation, as do the other microscopic variables. The integer part of this unusual dynamical variable can be regarded as the number of molecules of the system and the velocity of this number variable is interpreted as the tendency of the system to absorb or release a molecule with the material and energy reservoir. By employing the grand adiabatic ensemble,^{21,22} we may reexamine the free energy related functions and structural properties pertinent to open and semiopen systems. In this paper, we extend both the formal and practical aspects of the grand molecular dynamics (GMD) method. We chose water as an important case study. This provides us with a more stringent test of the method as the intermolecular forces are at least two orders of magnitude larger in this associated liquid than in the previous fluid Ar simulations.²² In addition, water at standard conditions is considerably more dense than the fluid Ar examples. As an accurate and precise evaluation of the free energy per particle for models of water and solutes in aqueous solution is of substantial current interest,^{1-3,25,26} we shall demonstrate the use of GMD to obtain such properties.

This paper is organized as follows: Section II describes the conceptual basis of implementing open systems for molecular dynamics simulations, the extended Lagrangian, and its related equations of motion. We also present a way to calculate the free energy, emphasizing the merits of letting the number of particles vary as a continuous dynamical vari-

^{a)} Present address: Molecular Simulations Inc., 199 S. Los Robles, # 540, Pasadena, CA 91101.

able. Section III briefly outlines the operational aspects of the GMD technique for polyatomic systems. In Sec. IV, we specify the model for our simulation and the parameters used. Specific results for our liquid water simulation with the new method are presented in Sec. V. Discussions concerning the physical implications and the possible further work conclude in Sec. VI. In the appendices, the calculation of the radial distribution function from open system dynamics trajectories and distinctive features of grand molecular dynamics addition and deletion steps are presented.

II. METHOD

In this section, we start from the augmented extended system Lagrangian and its related equations of motion¹²⁻¹⁴ and review the concept of a fractional number of particles.^{21,22} Then we further introduce the mathematics around which molecular open system dynamics is articulated. We will explore the physical meaning of the proposed simulations and study the relation to free energy and particle numbers. Finally, we discuss a simple way in which one can obtain the excess free energy through a series of simulations.

A. Equations of motion

In the present section, we first sketch the underlying Lagrangian and equations of motion for the grand adiabatic ensemble. More complete presentations are in the literature.^{21,22} We start by introducing an additional degree of freedom, the particle number N , as an augmentation variable into the equations of motion. The number of particles is made to be a continuous variable, denoted ν , and is written as a sum

$$\nu = N + \xi, \quad (1)$$

where N is the integer part of ν and ξ is the fractional remainder. Thus, $\nu(t)$ is the time-dependent number of particles of the system and includes the number of normal particles $N(t)$, which have normal interactions in the system, and $\xi(t)$, which represents a single fractional particle which interacts with the rest of the system as a normal particle except with weight ξ . In our previous applications of GMD to simple fluid systems, the extended potential energy coupling of the number of particles to its conjugate variable μ in the form μN proved difficult numerically and unnecessary formally. We found that other functions of the form $C(\mu)k(\nu)$, which proved more satisfactory in a numerical scheme, could be used without loss of formal generality. In association with the general function of ν , we thus define a thermodynamically conjugate parameter or function C . Specifically C is a function of relative free energy per particle. The function $C(\mu)k(\nu)$ is introduced below as an extended chemical potential function.

The complete set of equations of motion to be solved below represents an extended virtual system^{14,27} which can be transformed mathematically into the physical system. Because the extended virtual system is conservative, its Hamiltonian is a constant of motion, but the physical total energy of the system fluctuates as in an open system. All the equations of motion for $6N + 8$ degrees of freedom are solved without introducing a stochastic process throughout the en-

tire course of the simulations. The ansatz Lagrangian \mathcal{L} may be written for the molecular case as²²

$$\begin{aligned} \mathcal{L}_{N+\xi}(r^N, p^N, r_f, p_f, \nu, p_\nu) \\ = \sum_{i=1}^N \sum_{a=1}^M \frac{p_{ia}^2}{2m} + \sum_{i=1}^M \frac{p_{fi}^2}{2m} + \frac{p_\nu^2}{2W} \\ - \sum_{i=1}^{N-1} \sum_{a=1}^M \sum_{j>i}^M \sum_{b=1}^M U(r_{iab}^{oo}) \\ - (\nu - N) \sum_{i=1}^N \sum_{a=1}^M \sum_{j=1}^M U(r_{ij}^{of}) \\ - (\nu - N) \sum_{a=1}^{M-1} \sum_{b>a}^M U(r_{ab}^{ff}) + Ck(\nu), \end{aligned} \quad (2)$$

where M denotes the number of atoms per molecule, N represents the number of molecules in the system, and the upper labels o and f indicate normal full particle and fractional particle, respectively; p_ν is the general momentum associated with the extended degree freedom ν and W is its related mass parameter. Therefore, the first two terms in Eq. (2) give the total kinetic energy of the normal particles and the fractional particle, respectively; the kinetic energy with respect to the number variable ν is the third term; the fourth term represents the pairwise intermolecular potential between normal molecules i and j with atoms a and b ; the fifth term is the pairwise intermolecular potential between normal and fractional sites; the sixth term is the intramolecular potential for the fractional molecule; the last term is the coupling between the number of particles and the chemical potential written in a general notation. In previous work from this laboratory, it was demonstrated that this Lagrangian is consistent with the grand partition function for the physical system.²¹

In principle, the equations of motion for particle momentum as well as the position can be derived from any given Lagrangian of the system. Since the number of particles is treated as a continuous physical variable, its trajectory must also be determined by the equation of motion obtained from the extended Lagrangian \mathcal{L} and can, therefore, be acquired through numerical integration. The equations of motion corresponding to \mathcal{L} are

$$m\ddot{x}_{ia}^o = - \sum_{j=1}^N \frac{\partial U(r_{ij}^{oo})}{\partial r_{ij}} \frac{x_{ija}}{r_{ij}} - (\nu - N) \sum_{j=1}^M \frac{\partial U(r_{ij}^{of})}{\partial r_{ij}} \frac{x_{ija}}{r_{ij}}, \quad (3)$$

$$\begin{aligned} m\ddot{x}_{ia}^f = & - (\nu - N) \sum_{j=1}^N \frac{\partial U(r_{ij}^{fo})}{\partial r_{ij}} \frac{x_{ija}}{r_{ij}} - (\nu - N) \\ & \times \sum_{j=1}^M \frac{\partial U(r_{ij}^{ff})}{\partial r_{ij}} \frac{x_{ija}}{r_{ij}}, \end{aligned} \quad (4)$$

and

$$W\ddot{\nu} = C \frac{dk(\nu)}{d\nu} - \sum_{i=1}^N \sum_{j=1}^M U(r_{ij}^{of}) - \sum_{i=1}^{M-1} \sum_{j>i}^M U(r_{ij}^{ff}), \quad (5)$$

where a is a subscript representing the vector components in the x , y , and z directions. Therefore, \ddot{x}_{ia} is the acceleration of atom i along the direction a . The pair potentials $U(r_{ij}^{oo})$, $U(r_{ij}^{fo})$, and $U(r_{ij}^{ff})$ depend only on the magnitude of the sepa-

ration distance $r_{ij} = |r_i - r_j|$ between full atom i and full atom j , fractional atom i and full atom j , and fractional atom i and fractional atom j , respectively, where x_{ija} denotes the a component of the vector r_{ij} . Notice that $\nu - N$ provides a Born-like linear scaling of the associated terms, ranging between 0 and 1 and driven by Eq. (5) for the number of particles. There are a variety of ways to couple the fractional particle with the physical system and still preserve the grand adiabatic ensemble character of the corresponding dynamics. By choosing the term $Ck(\nu)$ (below) and a linear scaling of the potential, we have selected one such method. The formal proof demonstrating equivalence with the open system partition function follows Anderson's path¹² and has been given previously.²¹

$$f(r^N, p^N, r_f, p_f, \nu, p_\nu) = \frac{h^{-3N} \exp[-\beta H_{N+\xi}(r^N, p^N, r_f, p_f, \nu, p_\nu)]}{N! \Xi} \\ = \frac{h^{-3N} \exp[\beta Ck(\nu)] \exp[-\beta H_{N+\xi}^0(r^N, p^N, r_f, p_f, \nu, p_\nu)]}{N! \Xi}, \quad (6)$$

where $H_{N+\xi}(r^N, p^N, r_f, p_f, \nu, p_\nu)$ is the extended Hamiltonian associated with the Lagrangian \mathcal{L} in Eq. (2), which consists of the extended system free energy $Ck(\nu)$ and the extended virtual system energy is $H_{N+\xi}^0(r^N, p^N, r_f, p_f, \nu, p_\nu)$. In this case, we may write

$$\Xi = \sum_{N=0}^{\infty} \frac{h^{-3N}}{N!} \int_N^{N+1} \exp[\beta Ck(\nu)] d\nu \int \exp[-\beta H_{N+\xi}^0] \\ \times (r^N, p^N, r_f, p_f, \nu, p_\nu) dr^N dp^N dr_f dp_f dp_\nu \\ = \sum_{N=0}^{\infty} \int_N^{N+1} \exp[\beta Ck(\nu)] \int Q_{N+\xi}(V, T, p_\nu) d\nu dp_\nu, \quad (7)$$

which is an *extended grand partition function*. We define

$$Q_{N+\xi} = \frac{h^{-3N}}{N!} \int \exp[-\beta H_{N+\xi}^0] \\ \times (r^N, p^N, r_f, p_f, \nu, p_\nu) dr^N dp^N dr_f dp_f \quad (8)$$

as the *extended canonical partition function*. In this representation, the average Gibbs free energy may be rewritten in terms of the extended partition function

$$\langle G \rangle = \sum_{N=0}^{\infty} \int_N^{N+1} Ck(\nu) d\nu \int f(r^N, p^N, r_f, p_f, \nu, p_\nu) \\ \times dr^N dp^N dr_f dp_f dp_\nu \\ = \frac{1}{\Xi} \sum_{N=0}^{\infty} \frac{h^{-3N}}{N!} \int_N^{N+1} Ck(\nu) \exp[\beta Ck(\nu)] d\nu dp_\nu \\ \times \int \exp[-\beta H_{N+\xi}^0(r^N, p^N, r_f, p_f, \nu, p_\nu)] \\ \times dr^N dp^N dr_f dp_f \quad (9) \\ = \frac{1}{\Xi} \sum_{N=0}^{\infty} \int_N^{N+1} Ck(\nu) \exp[\beta Ck(\nu)] Q_{N+\xi} \\ \times (V, T, p_\nu) d\nu dp_\nu.$$

B. Gibbs energy in terms of the extended chemical potential function

In this section, we show how to connect the chemical potential with the specific scaling forms we have proposed in the previous section. Next, we will discuss a practical way to obtain the excess free energy per molecule.

For the purpose of this study, we restrict the discussion to open systems containing only one molecular species. By introducing a fractional molecule, the usual probabilities and distributions must also be extended.²¹ Thus, the ensemble probability density, which gives the probability both that the system contains N normal particles plus one fractional particle and that the coordinates and momenta of the particles are r^N, p^N, r_f, p_f, ν , and p_ν , is defined as

If we use the general function $Ck(\nu)$ (see below), then we have another representation of Gibbs energy in terms of the thermodynamic parameters

$$\langle G \rangle = \langle \mu N \rangle = \langle Ck(\nu) \rangle. \quad (10)$$

We will use the different representations of the Gibbs energy in Eqs. (9) and (10) to study the connection between the chemical potential μ and the extended chemical potential function $Ck(\nu)$ that we propose. Equation (9) can be simplified by interchanging the derivative with respect to C with the summation. For the equation for the Gibbs free energy at constant temperature, this leads to

$$\langle G \rangle = \frac{C}{\Xi} \sum_{N=0}^{\infty} \int_N^{N+1} \frac{1}{\beta} \frac{\partial \exp[\beta Ck(\nu)]}{\partial C} \\ \times Q_{N+\xi}(V, T, p_\nu) d\nu dp_\nu \\ = \frac{Ck_B T}{\Xi} \frac{\partial}{\partial C} \left\{ \sum_{N=0}^{\infty} \int_N^{N+1} \exp[\beta Ck(\nu)] \right. \\ \left. \times Q_{N+\xi}(V, T, p_\nu) d\nu dp_\nu \right\} \quad (11) \\ = Ck_B T \frac{\partial \ln \Xi}{\partial C}.$$

For an infinitesimal change in the parameter C under isothermal and constant volume conditions, the grand thermodynamic potential $-PV$ will vary as

$$-k_B T \frac{\partial \ln \Xi[\mu(C), V, T]}{\partial C} \Big|_{T, V} = -V \frac{\partial P}{\partial C}, \quad (12)$$

where we have written μ the chemical potential as a function of C the numerical parameter from Eq. (2). Substituting Eq. (12) into Eq. (11), we obtain the free energy

$$\langle G \rangle = CV \frac{\partial P}{\partial C}. \quad (13)$$

Furthermore, at constant volume and temperature¹⁷

$$VdP = Nd\mu. \quad (14)$$

Comparing Eqs. (13) and (14), the relation of the chemical potential to the parameter C can be deduced as

$$\mu = \frac{G}{N} = \frac{CV}{N} \frac{\partial P}{\partial C} = \frac{CV}{N} \frac{N}{V} \frac{\partial \mu}{\partial C} = C \frac{\partial \mu}{\partial C}, \quad (15)$$

which results in a simple first-order linear differential equation. This demonstrates that μ and C are simply related to each other and this serves to define C as a function of the free energy of the system. In this case, we characterize the open system by choosing the constant C . We have a simple solution

$$\mu = A(V, T)C, \quad (16)$$

where $A(V, T)$ is an integration constant which is independent of C and is only a function of the volume V and temperature T of the system. This integration constant can be calculated as an ensemble average in the simulations.

For the grand canonical ensemble, given the probability $P(N)$ that the system contains precisely N particles, we have

$$\begin{aligned} \mu &= -k_B T \Delta \ln Q_N(V, T)|_{N=\bar{N}} \\ &= -k_B T (\ln Q_{N+1} - \ln Q_N)|_{N=\bar{N}} \end{aligned} \quad (17)$$

at the average particle number value $\bar{N} = N_0$, or the system's most probable particle number value.

By analogy,¹⁷ the probability $P(\nu)$ for our open systems can be obtained by integrating $f(r^N, p^N, r_f, p_f, \nu, p_\nu)$ over the new phase space generated by the Lagrangian of Eq. (2). That is,

$$\begin{aligned} P(\nu) &= \int \int f(r^N, p^N, r_f, p_f, \nu, p_\nu) dr^N dp^N dr_f dp_f dp_\nu \\ &= \frac{1}{\Xi} \exp[\beta C k(\nu)] \int Q_\nu(V, T, p_\nu) dp_\nu \end{aligned} \quad (18)$$

and the most probable number state will be governed by the condition

$$\frac{\partial \ln P(\nu)}{\partial \nu} = \frac{\partial \ln z}{\partial \nu} + \frac{\partial \ln Y_\nu}{\partial \nu} = 0, \quad (19)$$

where $z = \exp[\beta C k(\nu)]$ is the *extended activity* and Y_ν is integral of Q_ν along the particle number momenta space. That is,

$$\begin{aligned} Y_\nu &= \int Q_\nu(V, T, p_\nu) dp_\nu \\ &= Q_{N+\xi}(V, T) \sqrt{\frac{2\pi W}{\beta h^2}}, \end{aligned} \quad (20)$$

where

$$\begin{aligned} Q_{N+\xi}(V, T) &= \frac{h^{-3N}}{N!} \int \int \exp[-\beta H'_{N+\xi} \\ &\quad \times (r^N, p^N, r_f, p_f)] dr^N dp^N dr_f dp_f \end{aligned} \quad (21)$$

and the Hamiltonian $H'_{N+\xi}$ is the total kinetic and the total potential energies only from the normal and fractional particles. Thus, we find

$$\begin{aligned} \beta C \frac{\partial k(\nu)}{\partial \nu} \Big|_{\nu=\bar{N}} &= - \frac{\partial \ln Y_\nu}{\partial \nu} \Big|_{\nu=\bar{N}} \\ &= - \frac{\partial \ln Q_\nu(V, T)}{\partial \nu} \Big|_{\nu=\bar{N}}. \end{aligned} \quad (22)$$

Since Y_ν is the product of $Q_\nu(V, T)$ and W , the thermodynamic derivatives of $\ln Y_\nu$ should be independent of this general mass W . In principle, the mass W does not affect the average number of particles or other thermal averages. If one substitutes $\partial \nu = 1$ and $\nu = \bar{N}$, then we have

$$\begin{aligned} \beta C [k(\nu+1) - k(\nu)]|_{\nu=\bar{N}} \\ = -(\ln Q_{N+1} - \ln Q_N)|_{N=\bar{N}}. \end{aligned} \quad (23)$$

Comparing Eqs. (17) and (23), we have the chemical potential as

$$\mu = C [k(\nu+1) - k(\nu)]|_{\nu=\bar{N}}. \quad (24)$$

Comparing Eq. (16) with Eq. (24), the integration constant $A(V, T)$ will simply be equal to $k(\nu+1) - k(\nu)$ when ν is equal to the average number of particles \bar{N} for a large number of particles. Since there are many ways to couple the particle with the system, we can choose the derivative of the scaling function $k(\nu)$ to be equal to 1 at $\nu = \bar{N}$; then the chemical potential will simply be equal to the parameter C we have chosen.

Using a time average, the chemical potential defined by Eq. (15), with the conditions mentioned in the preceding paragraph, can be calculated as

$$\mu = \frac{\langle Ck(\nu) \rangle}{\langle \nu \rangle}. \quad (25)$$

This argument shows that a wide range of coupling functions may be used to fix the chemical potential without loss of generality.

III. IMPLEMENTATION OF OPEN SYSTEM DYNAMICS

In the process of implementing GMD, there are some technical features and conceptual problems which do not occur in other ensembles. In this section, solutions to those problems and our reasoning are presented. The integration of the equations of motion (3)–(5) can be performed by a variety of methods. In this work, we shall report the results of implementing the velocity Verlet algorithm.¹⁷ We have also, in previous work,^{21,22} used a six value Gear predictor corrector for the numerical integration. As our potential function for the number variable [Eq. (5)], we use the same general number function employed previously

$$k(\nu) = \nu e^{(\nu/\nu_0)^2 - 1}, \quad (26)$$

where ν_0 is a constant chosen for stability of the equations of motion.

Given the introduction of the principle of a fractional particle to overcome the discontinuity of the particle number N ,^{21,22} it becomes essential to deal with the deletion and insertion of an infinitesimal fractional molecule as the result of the exchanges between the bath and the system. The procedure for insertion or deletion of the fractional particle is the counterpart of particles exchanging with the external bath. For GMD, the standard molecular dynamics algorithms for

constant N must be supplemented by algorithms for inserting and deleting particles. The deletion of a fractional particle occurs when ξ becomes 0, at which point another particle may be chosen for deletion or a spot determined for insertion, depending on the velocity of the number variable \dot{v} . Similarly, a fractional particle becomes a full particle when ξ becomes 1. The choice of position for the operation, i.e., where to insert the next particle or which particle to next delete, is based on the work required to insert or delete a particle from a given point. The estimate is based on the value of the potential energy content of the previous fractional particle. To insert, we use a grid search for finding a likely cavity and orientation. To delete, we find the particle in a region of poor chemical potential with a potential (fully interacting) closest to the last fractional particle. More details are given in Appendices B and C.

IV. MODEL AND PARAMETERS

We started the simulations with a system of 216 SPC model (simple point charge) water molecules equilibrated previously in the microcanonical (N, V, E) ensemble.^{28,29} Their bonds were constrained with the RATTLE algorithm.³⁰ Periodic boundary conditions on a cubic box with edge length 18.62 Å were used. The initial set of velocities of the system corresponded to a temperature of 300 K. A continuous switching function for the pair interactions between 7.8–8.8 Å was employed. We used a time step of 2.0 fs and equilibrated each run of the system for 20.0 ps before data collection. The results of several simulations are analyzed below.

Before we start the open system simulation, there is a choice that has to be made corresponding to the fictitious mass W used for the additional degree of freedom v in Eq. (2). Previously, several masses were considered for liquid argon. However, Eq. (5) shows that because of the dependence of the number acceleration on potential energy, a different mass dependence is expected for water. We tested three different masses—5.00, 15.00, and 50.00 kcal ps²/mol. We set the parameter C to be -2.67 kcal/mol and, in each case, it yielded an average number of particles near 218.

The general mass parameter W given in units of $\text{energy} \times \text{time}^2$ does not enter into definition of any measurable equilibrium thermodynamical property of the physical system, just as the other physical masses involved in the system. The direct effect of W is to govern the time scale involved in the interaction of the system with the material reservoir. The “lighter” W , the faster the particles are added, or in terms of Gibbs replicas, the more porous the walls around the system. In the large W limit, this system approaches the closed system behavior, i.e., the system is being detached from the material reservoir via less porous walls. In the regime of very low mass parameter values, numerical inaccuracies arise analogous to those observed for the masslike parameter for the Nosé thermostat,¹⁴ or for the masslike parameter in constant pressure/stress dynamics.¹² In the harmonic approximation, the period of the fluctuations in number variable is proportional to the square root of W . Very small values give shorter periods for the number fluctuations causing too frequent insertion and deletion steps,

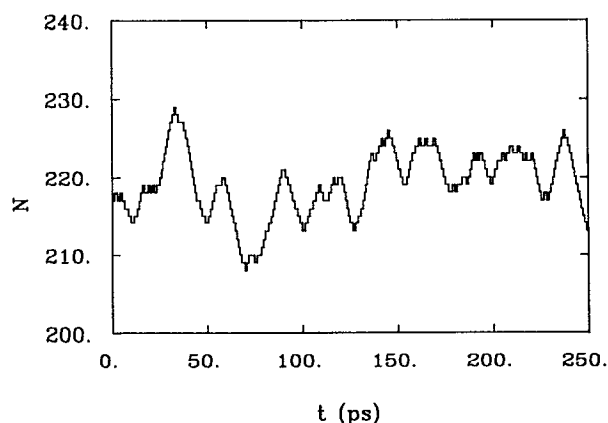


FIG. 1. The number of water molecules vs the simulation time at an average temperature of 300 K with the mass parameter $W = 5.00$. The average number of particles is 218.

which in turn causes numerical inaccuracies in the conservation of the extended Hamiltonian. Very large masses lead to instabilities similar to those seen with Nosé's thermostat.¹⁴

V. RESULTS

First we shall present results of the time-dependent number fluctuations with respect to mass. As shown in the Figs. 1, 2, and 3, the simulation with the smallest mass $W = 5.00$ releases or absorbs particles from the bath the most rapidly. The average time during which the system stays in one particle number state and does not exchange with the outside bath is about 0.5 ps for this mass parameter.

For the largest mass $W = 50.00$, the simulation stays much longer in each particle number state, on the average about 3.9 ps. The large mass results in a slow time evolution of the system with respect to the number and its fluctuations. The system has to go far beyond the average position to gain strong enough repulsion to turn back [see Eq. (5)]. The middle simulation with $W = 15.00$ gives the system an average of 1.8 ps in each number state. In Table I, we list a few ensemble averages collected in the simulations. We can see they are nearly identical for different masses, as expected from the analysis presented above. The rest of our studies

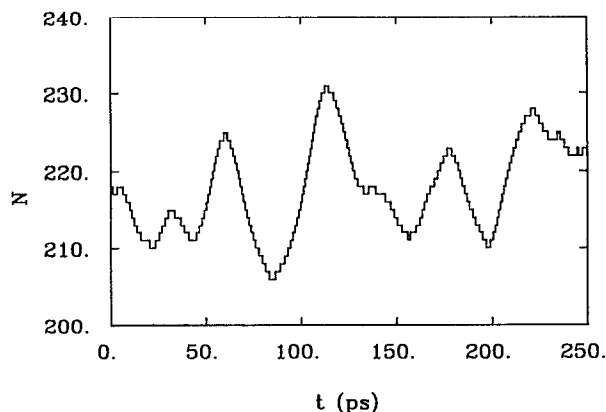
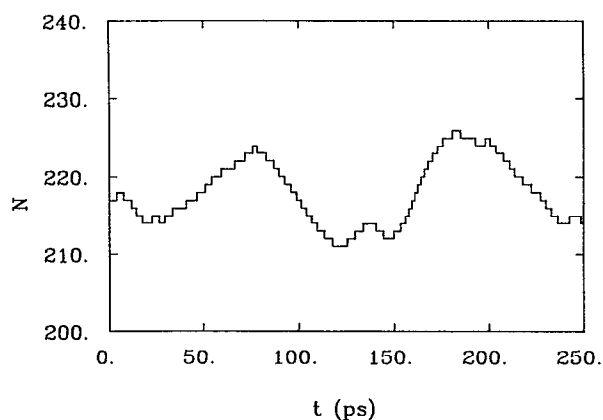


FIG. 2. The same as above, but with mass parameter $W = 15.00$.

FIG. 3. The same as above, but with mass parameter $W = 50.00$.

were carried out with a mass of 15.00.

In comparison with the simulation done in liquid Ar, where the mass W was set to unity in reduced units or 1.05 kcal ps²/mol, the water model requires a heavier mass.²² From Eq. (5), the acceleration of the particle exchange with the external bath depends on the interaction potential between the fractional particle and the normal particle. The total contribution from the interaction potential, along with the physical mass contribution, is at least an order of magnitude larger than liquid Ar. Table II also presents a few comparisons of the thermodynamic (energetic) properties of the SPC water model for several open systems. The mean interaction energy per molecule in the liquid state obtained from the GMD simulation is consistent with the one from a closed system.^{6,25,28,31}

Figures 4 and 5 give the average potential $\langle U \rangle$ and kinetic $\langle T \rangle$ energies, respectively. The expected anticorrelation is observed. The correlation of $\langle T \rangle$ with $N(t)$ is reasonable given the derivation of temperature for the grand adiabatic ensemble.²¹ By including the contribution from $\mu(\nu)$, we find the Hill energy $L = E - \mu N$ is conserved to better than one part in 10^4 kcal/mol for the system containing an average number of 218 water molecules.

The radial distribution functions $g_{OO}(r)$, $g_{OH}(r)$, and $g_{HH}(r)$ for open and closed water systems of equal density show deviations smaller than can be resolved graphically by direct comparison. This is expected on fundamental grounds. As these are well known, we have chosen not to display them. Their calculation is outlined in Appendix A. The time-dependent velocity autocorrelations show similar quality behavior, as expected.²²

Figure 6 shows the mean-squared displacement of the

TABLE I. A comparison of different masses.

W	50.00	15.00	5.00
$\langle N \rangle$	217.9	217.7	217.8
μ (kcal/mol)	-5.43	-5.40	-5.42
E (kcal/mol)	-10.18	-10.21	-10.19

TABLE II. A comparison of energies.^a

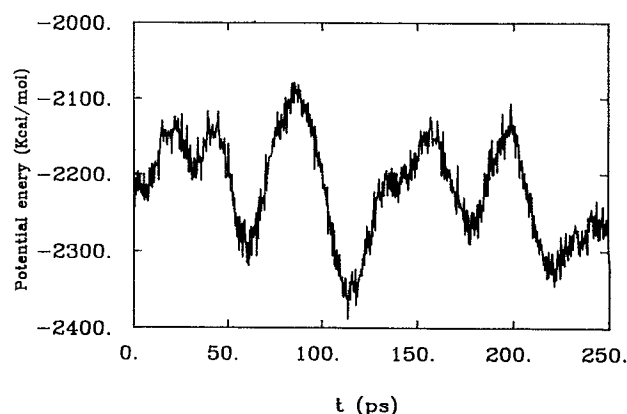
$\langle N \rangle$	$\langle T \rangle$	$\langle U \rangle$	$\langle E \rangle$
232.9	1.77 ± 0.06	-10.20 ± 0.15	-8.43 ± 0.11
226.5	1.79 ± 0.07	-10.18 ± 0.20	-8.40 ± 0.16
221.3	1.77 ± 0.08	-10.19 ± 0.23	-8.42 ± 0.17
219.9	1.77 ± 0.07	-10.19 ± 0.17	-8.42 ± 0.13
217.7	1.75 ± 0.07	-10.18 ± 0.17	-8.42 ± 0.13
215.3	1.75 ± 0.08	-10.19 ± 0.25	-8.44 ± 0.19
209.7	1.72 ± 0.07	-10.19 ± 0.23	-8.47 ± 0.17

^a Units of kcal/mol.

open system vs the closed system with respect to time, the slope of which gives the diffusion constant. We can see that the diffusion in the open system is essentially the same as the closed system. In previous work,³² the diffusion constant obtained by computer simulation was found to be very sensitive to the model and method used in the simulation. Apparently much larger differences than displayed in Fig. 6 can be generated by common changes in boundary conditions and interaction cutoffs.

By testing different values of the parameter C (or μ), we can explore the system's behavior with respect to a reservoir with a higher or lower chemical potential. In analogy to an earlier experiment on a simple fluid, we show the nonequilibrium response of the system to a slightly lower chemical potential. Figure 7 shows the system will be driven quickly to a new equilibrium for C equal to -2.07 kcal/mol, given an initial configuration obtained from C value set to -2.67 . This corresponds to changing the chemical potential by less than 0.25 kcal/mol.

An important feature of the open system approach is the ability to obtain the average number given the input chemical potential. As we stated earlier in Eq. (24), the excess free energy per particle can be evaluated through calculating the difference of the function $k(\nu)$ at the average particle number. We simulated 400 ps for each of several different values of C (from -2.07 to -3.07 kcal/mol) and obtained the corresponding average particle numbers (from 232.9 to 209.7 per box). By fitting these numbers to a straight line

FIG. 4. The potential energy of the open system vs the simulation time at a temperature of 300 K with mass parameter $W = 15.00$.

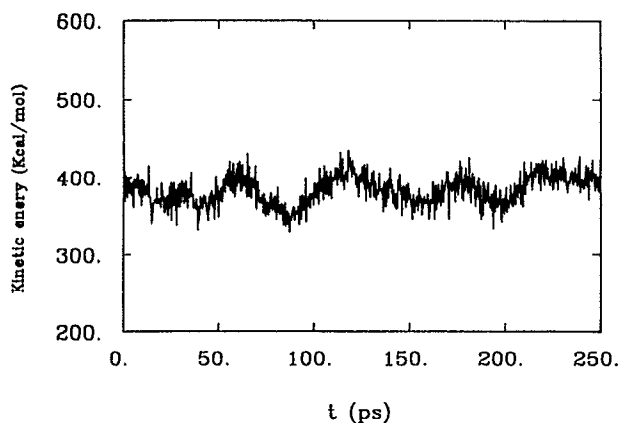


FIG. 5. The kinetic energy of the open system vs the simulation time at a temperature of 300 K with mass parameter $W = 15.00$.

$$\mu^{\text{ex}} = a_1 \bar{N} + b_1, \quad \text{or} \quad C = a_2 \bar{N} + b_2,$$

we have $a_1 = 0.01411$ and $b_1 = -8.500$, $a_2 = 0.04468$, and $b_2 = -12.42$. Table III lists the values of C , the average number of molecules and excess chemical potential per run.

Figures 8 and 9 display the relation between the special chemical potential C vs the average number of molecules and the excess free energy vs the average number of molecules, respectively. As shown earlier, the relation between C and μ is linear. From Fig. 10, we obtain the chemical potential at nearby liquid densities by a simple linear extrapolation. By interpolating to $\bar{N} = 216$, which corresponds to a density of 1.0006 g/cm^3 , we find an excess free energy per particle of -5.45 kcal/mol . This is comparable to the result of Hermans and co-workers³³ who, using thermodynamic integration, obtained a free energy per particle of -5.6 kcal/mol in a system of 80 SPC water molecules with a 6 \AA cutoff and density of 0.995 at 300 K . By interpolating to their density, we obtain a value of -5.47 kcal/mol . Those authors also tested the effects of different cut-off distances. For the same system with a 10 \AA cutoff, the excess free energy found was

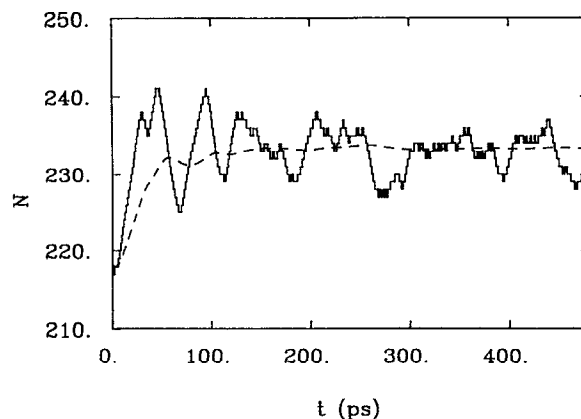


FIG. 7. The number of water molecules when $C = -2.07 \text{ kcal/mol}$ vs time at a temperature of 300 K with the initial coordinate obtained from $C = -2.67 \text{ kcal/mol}$; the dashed line is the cumulative time average for the number of water molecules.

-5.8 kcal/mol , although the density of 1.017 g/cm^3 was slightly different. Thus, the change in cutoff affects the result of the chemical potential calculation similarly to a density change of less than 1% . The free energies of different types of water models were also calculated.³³ For SPCE, TIP3P, and TIP4P, the results were -5.5 , -5.4 , and -5.3 kcal/mol , respectively. Mezei,²⁶ in an early free energy calculation, reported that the chemical potential for 64 SPC molecules at 25°C and density 1 g/cm^3 is -5.64 kcal/mol , which is comparable to the numbers of Hermans *et al.*³³ but about 0.2 kcal/mol different from our value. Using a grand canonical Monte Carlo procedure, Mezei found -5.47 kcal/mol in excellent agreement with our findings.³⁴ Finally, Sarkisov and co-workers,³⁵ using a canonical Monte Carlo method, observed the excess free energy to be -5.71 kcal/mol for 64 water molecules with a density of 0.995 g/cm^3 . There appears to be a small difference that may be ensemble dependent, but we have made no systematic study to confirm this.

We also considered the analysis of the set of simulations for different values of C vs the total simulation time used for the analysis. In Fig. 10, we compare those results obtained with the entire set of 400 ps simulations and shorter windows. It is apparent that one can use our method with shorter time series (e.g., 60 ps) to predict the model's chemical potential at a given density with reasonable chemical accuracy ($\sim 0.1 \text{ kcal/mol}$). This shows the potential of the method for making efficient evaluations of the chemical po-

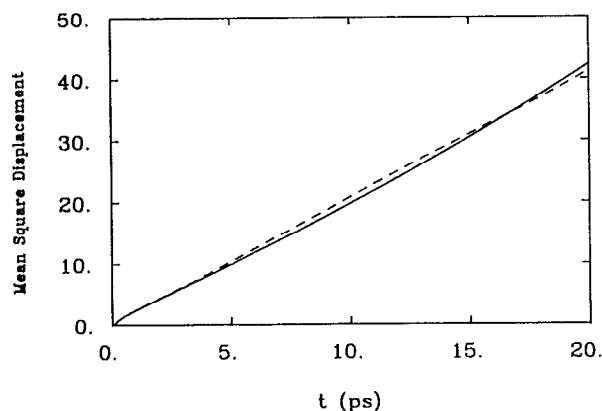


FIG. 6. The comparison of the mean-squared displacement vs time calculated for closed and open systems using the center-of-mass coordinates of the water molecules. The solid line indicates the open system and the dashed line is the closed system.

TABLE III. A comparison for different C values.

$C(\text{kcal/mol})$	$\langle N \rangle$	$\mu^{\text{ex}} (\text{kcal/mol})$
-2.07	232.9	-5.32
-2.27	226.5	-5.22
-2.47	221.3	-5.23
-2.57	219.9	-5.37
-2.67	217.7	-5.40
-2.87	215.3	-5.67
-3.07	209.7	-5.52

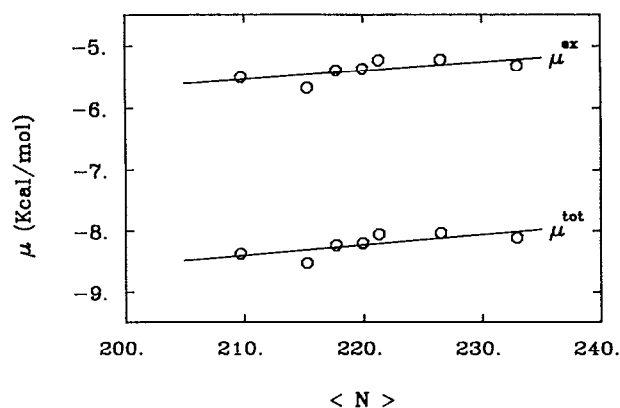


FIG. 8. The excess free energies per molecule μ^{ex} (kcal/mol) and total free energies per molecule μ vs the average number of water molecules. The circles are obtained from the simulation and the solid line is fitted.

tential over a continuous range of densities, compared to the currently popular thermodynamic integration¹⁹ or the use of perturbation theory,²⁰ where similar quality results for only one density are obtained only beyond a half-nanosecond of simulated time. We mention at this point that these grand molecular dynamics simulations take only about 1% more time than a conventional NVE simulation, where $N_{\text{NVE}} = \langle N \rangle_{\mu V L}$ and the same interaction truncation scheme is used.

VI. CONCLUSION

The use of the Born-charging concept of a fractional particle which enables a practical coupling of the system with a particle bath captures the essence of the grand adiabatic ensemble. By extending the concept of the particle number from integer to a real variable and providing the equations of motion, we have presented in this paper not only a complete method of modeling grand ensembles, but also a practical treatment applied to an ensemble of a polyatomic system. This method is, in essence, similar to other members of the class of Anderson-like computational approaches to

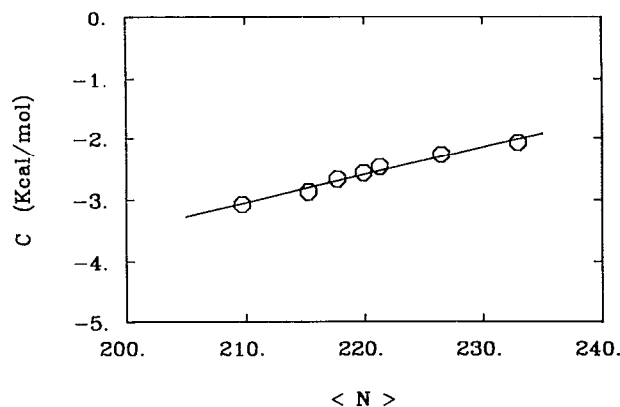


FIG. 9. The parameter C (kcal/mol) vs the average number of water molecules in the simulation box. The circles are obtained from the simulation and the solid line is fitted.

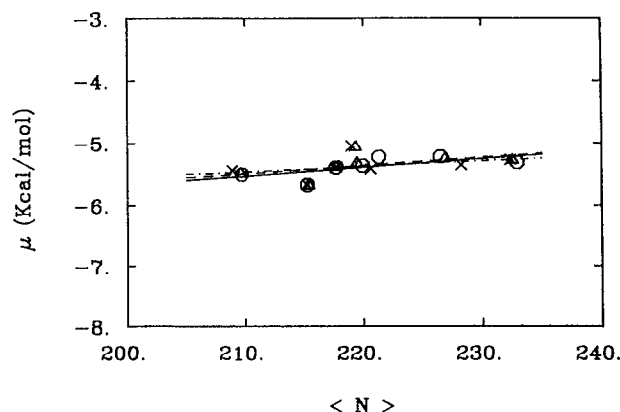


FIG. 10. The excess free energies μ^{ex} (kcal/mol) vs the average number of water molecules. Symbols are measured and lines are fitted. The solid line and circles are data averaged over 300 ps of the simulations. The dashed line and \times 's are data averaged from the last 120 ps of the simulations. The dotted line and triangles are data averaged from the last 60 ps of the simulations.

molecular dynamics simulations in various ensembles,¹² except here the particle number is treated as a dynamic variable. Although the force field and implementations we used for our system were certainly not unique, this first dynamic simulation attempt of a polyatomic open system demonstrated several powerful features of the method. In addition, we have shown that this technique gives a very simple way to predict directly the free energy per particle for the physical system over a range of densities.

In our simulation, we presented thermodynamic comparisons, including the mean interaction energy per water molecule, which match those obtained from the equivalent closed system in earlier studies. We found the radial distribution functions which summarize the thermodynamic behavior of the liquid for the open system can be superimposed on those obtained for the corresponding closed system.^{28,31} The results from our GMD simulation of the free energy per particle as a function of the average number of particles in a given volume are in very good agreement with the predictions made in earlier studies obtained through Monte Carlo, thermodynamic integration, and perturbation method studies.^{3,25,26} The use of this method in liquid water, demonstrates that it is capable of yielding reasonable results in systems with substantially higher density and interaction strengths. Work is currently in progress on ionic solutions and phase equilibria.

ACKNOWLEDGMENTS

The authors thank the Texas Coordinating Board through the Advanced Research Program and the Robert A. Welch Foundation for partial support of this work. IBM is acknowledged for a generous grant of 3090 time. B.M.P. was an Alfred P. Sloan Fellow for 1989–91.

APPENDIX A: CALCULATION OF THE RADIAL DISTRIBUTION $g(r)$

The calculations for many physical observables of the open system are different operationally from those systems

with a fixed number of particles. Many of the important structural characteristics of a liquid are described conveniently by the radial distribution function (RDF). The pure water radial distribution functions $g_{OO}(r)$, $g_{OH}(r)$, and $g_{HH}(r)$ have been well studied in the closed system case.^{28,29,31} The quantity $4\pi\rho r^2 g_{ab}(r) dr$ is the probability of finding a pair of atoms at distance r relative to the probability expected for a completely random distribution at the same density. Since the density ρ_0 is no longer constant in the grand canonical ensemble, the calculation must be modified slightly.^{17,23}

If we view the system as a canonical ensemble at every instant of time, which subsequently is to be assembled into the grand canonical ensemble, then the radial distribution function should be just the collection of the radial distributions at each different canonical ensemble (value of N) and is defined conveniently as

$$g_{ab}(r) = \sum_{N=0}^{\infty} \rho_N^{-2} \left\langle \sum_i \sum_{j>i} \delta(r - |\mathbf{r}_i - \mathbf{r}_j|) \right\rangle_{\text{canonical}}, \quad (27)$$

or

$$g_{ab}(r) = \left\langle \frac{n_{ab}(r)}{\rho_v^2} \right\rangle_{\text{grand canonical}}, \quad (28)$$

where $\rho = \nu/V$ and ν is the continuous number of particles at every step and the summation of all the pairs will include the fractional particle as well as normal particles.

APPENDIX B: DELETION

We chose to delete the fractional particle according a criterion similar to creation, i.e., based on the fractional particle number value ξ . When this value decreases to near zero during a time step, deletion takes place. Then we delete the fractional particle with its coordinates, momenta, and interactions with other particles. If the number velocity ν remains negative, we must choose another candidate for deletion from among the normal particles. One would imagine that the most favorable removal would be from the densest part of the system or a particle with high potential energy. To aid in the numerical continuity of Eq. (5), we pick a molecule which has the closest potential energy to a complete molecule at the deleted fractional particles' coordinates. This is usually a molecule near the highest potential energy. We then consider this molecule to be the new fractional particle.

When the particle number value ν is very close to the next lower integer from the equation of motion [Eq. (4)], the repulsive force on the fractional particle due to the contributions of the normal particles, which is proportional to the value of ξ , could be very small. Accordingly, the fractional particle could penetrate very deeply into a normal particle. Examination of the equation of motion for ν [Eq. (5)] shows that the number acceleration can become negative and large. As a consequence of finite numerical time steps, precision can be maintained by reducing the time step or by using an algorithm designed for multiple time scales. In order to avoid this problem with a single time step, we adopt a cutoff such that the deletion is considered complete when ξ is between 0.01 and 0.001. As the contribution of such a small fractional particle to the system's free energy is essentially

zero, the procedure is justified numerically. To compensate for the very small changes in the kinetic and potential energies due to this truncation, we can scale the total kinetic energy after every deletion. Even without such scaling, the energy L is still conserved to nearly a part in 10^3 kcal/mol.

APPENDIX C: ADDITION

The addition procedure, which is different from the previous case, requires additional attention. While deletion involves a restricted number of particles, addition requires an inspection of the entire space to choose the most appropriate location.

When ν increases to the next integer, the former fractional particle emerges as a full particle. Concurrently, we are compelled to find a set of coordinates for another new fractional particle which has the closest potential energy with respect to the former fractional particle because of the reason stated earlier. In this case, we shall search for a region of space with very few particles.

To examine the entire space in order to find the proper position for the new fractional particle, we employ a grid method to slice the box used in simulation into smaller boxes. Experience has shown that a convenient grid is generated when the number of these small boxes is the same as or a little larger than the number of the molecules in the system, so that there is nearly always at least one small box which is not occupied. After finding the empty (or least dense) boxes, we place a new fractional molecule in the center of each trial box. This is similar to the approaches which were used in the cavity-biased (μVT) Monte Carlo¹ and excluded volume-biased test particle methods for obtaining chemical potential of simple fluids.² Since the water molecule is dipolar, orientation must also be taken into account. Therefore, for each of the boxes, we calculate separately the total potential energies between the normal water molecules and the fractional water molecule with oxygen held at the center of the small box for 200 possible orientations previously stored. Then, out of these 200 configurations, the one which has the closest larger potential compared to the formal fractional particle is elected. Afterward, we apply briefly a steepest descent energy minimization procedure for approximately 10 steps to relax this water molecule to the desired energy. If the potential of this molecule is close (0.001 kcal) to the one of the former fractional molecule, we consider the new fractional particle to be found. Otherwise, we continue to apply the same procedures on each of the small boxes. If the condition is not satisfied, we perform the energy minimization procedure on the test molecule for up to 100 steps. Then we consider this test molecule as the new fractional particle. The numerical error due to the approximated location chosen for the fractional particle can disturb the system. When the insertion of an infinitesimal fractional particle into the system occurs, the velocity of this fractional particle is assigned to be zero. The reason is that the mass of the fractional particle is formally infinitesimally small when the insertion takes place. Essentially, it corresponds to an ideal gas particle. Once it obtains a finite interaction with the bath, it also gains temperature from the bath.

- ¹ M. Mezei, *Mol. Phys.* **40**, 901 (1980).
- ² G. L. Deitrick, L. E. Scriven, and H. T. Davis, *J. Chem. Phys.* **90**, 2370 (1990).
- ³ M. Mezei, S. Swaminathan, and D. L. Beveridge, *J. Am. Chem. Soc.* **100**, 3255 (1978).
- ⁴ R. W. Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954).
- ⁵ T. P. Straatsma, Ph.D. thesis, University of Groningen, 1987.
- ⁶ T. P. Straatsma, H. J. C. Berendsen, and J. P. M. Postma, *J. Chem. Phys.* **85**, 6720 (1986).
- ⁷ M. Mezei, *J. Chem. Phys.* **86**, 7084 (1987).
- ⁸ W. L. Jorgensen and C. Ravimohan, *J. Chem. Phys.* **83**, 3050 (1985).
- ⁹ D. J. Tobias and C. L. Brooks, *Chem. Phys. Lett.* **156**, 256 (1989).
- ¹⁰ B. J. Alder and T. A. Wainwright, *J. Chem. Phys.* **27**, 1208 (1957); **31**, 459 (1959).
- ¹¹ A. Rahman, *Phys. Rev. A* **136**, 405 (1964).
- ¹² H. C. Andersen, *J. Chem. Phys.* **72**, 2384 (1980).
- ¹³ M. Parrinello and A. Rahman, *Phys. Rev. Lett.* **45**, 1196 (1980).
- ¹⁴ S. Nosé and M. L. Klein, *Mol. Phys.* **50**, 1055 (1983); S. Nosé, *ibid.* **52**, 255 (1984).
- ¹⁵ Z. W. Salsburg, J. D. Jacobson, W. Fickett, and W. W. Wood, *J. Chem. Phys.* **30**, 67 (1959).
- ¹⁶ J. P. Valleau, and G. M. Torrie, in *Statistical Mechanics A. Modern Theoretical Chemistry*, edited by B. J. Berne (Plenum, New York, 1976), Vol. 5, p. 169.
- ¹⁷ M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University, New York, 1987).
- ¹⁸ C. F. Anderson and T. Record. *Annu. Rev. Biophysics Biophys. Chem.* **19**, 301 (1990).
- ¹⁹ M. Mazor and B. M. Pettitt, *Mol. Simul.* **6**, 1 (1991).
- ²⁰ D. A. Pearlman and P. A. Kollman, *J. Chem. Phys.* **94**, 4532 (1991); M. Mitchell and J. A. McCammon, *J. Comp. Chem.* **12**, 271 (1991).
- ²¹ T. Cagin and B. M. Pettitt, *Mol. Simul.* **6**, 5 (1991).
- ²² T. Cagin and B. M. Pettitt, *Mol. Phys.* **72**, 169 (1991).
- ²³ J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Wiley, New York, 1986), p. 26.
- ²⁴ D. A. McQuarrie, *Statistical Mechanics* (Harper & Row, New York, 1976).
- ²⁵ J. Hermans, A. Pathiaseril, and A. Anderson, *J. Chem. Soc.* **110**, 5982 (1988).
- ²⁶ M. Mezei, *Mol. Phys.* **47**, 1307 (1982); **67**, 1205 (1989) (erratum).
- ²⁷ J. Jellinek, *Phys. Rev. A* **38**, 3069 (1988).
- ²⁸ L. X. Dang and B. M. Pettitt, *J. Chem. Phys.* **91**, 3349 (1987).
- ²⁹ H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, 1981).
- ³⁰ H. C. Andersen, *J. Chem. Phys.* **52**, 24 (1983).
- ³¹ F. H. Stillinger and A. Rahman, *J. Chem. Phys.* **60**, 1545 (1974).
- ³² P. E. Smith and B. M. Pettitt, *J. Am. Chem. Soc.* **113**, 6029 (1991).
- ³³ J. Hermans, A. Pathiaseril, and A. Anderson, *J. Am. Chem. Soc.* **110**, 5982 (1988).
- ³⁴ M. Mezei, *Mol. Phys.* **67**, 1207 (1989).
- ³⁵ G. N. Sarkisov, V. G. Dashevsky, and G. G. Malenkov, *Mol. Phys.* **27**, 1249 (1974).